REPORT DOCUMENTATION PAGE

AFRL-SR-AR-TR-03-

0471

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
05-11-2003	Final	01/02/00 - 30/11/02

4. TITLE AND SUBTITLE
Large-Scale Optimization Methods with a Focus on Chemistry Problems

5a. CONTRACT NUMBER
F49620-00-1-0162

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Schnabel, Robert B.
Byrd, Richard. H.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Colorado at Boulder
Office of Contracts & Grants
3100 Marine St., Rm. 481, 572 UCB
Boulder, CO 80309

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research
801 N. Randolph St., Rm. 732
Arlington, VA 22203-1977

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

20031121 010

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Our research under this grant has developed general large-scale global optimization techniques in the context of solving protein structure prediction problems. Over the course of the research period the thrust of the research was primarily in two areas. The first area provided advances in our techniques for handling very large, complex proteins of arbitrary secondary and tertiary structure. The second area focused on advances in the efficiency of the algorithm on large-scale computing platforms. These advancements enabled us to significantly improve our effort in the CASP community-wide competition for protein structure prediction. During the CASP5 competition in summer of 2002, we attempted blind predictions of 20 proteins with sizes and structures far more complex than any previously attempted by our group. The results of assessment for the category of proteins which contain ``new folds" (i.e. proteins which were not known by existing databases) show that our group's predictions were in the top 10% of groups which predicted these targets. Much of this research is joint work with the bio-chemistry group headed by Dr. Teresa Head-Gordon at UC Berkeley, and the computer science group of Dr. Silvia Crivelli at Lawrence Berkeley Laboratory.

15. SUBJECT TERMS
Global optimization, molecular chemistry, large-scale optimization, protein folding

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Schnabel, Robert B.
Unclassified	Unclassified	Unclassified		10	19b. TELEPHONE NUMBER *(include area code)* 303-492-7554

Final Technical Report

U.S. Air Force Grant F49620-00-1-0162

Large Scale Optimization Methods with a Focus on Chemistry Problems

Richard Byrd, Robert Schnabel
Department of Computer Science
University of Colorado, Boulder
Boulder, CO 80309-0430

1 February 2000 - 30 November 2002

1 Abstract

Our research under this grant has developed general large-scale global optimization techniques in the context of solving protein structure prediction problems. Over the course of the research period the thrust of the research was primarily in two areas. The first area provided advances in our techniques for handling very large, complex proteins of arbitrary secondary and tertiary structure. The second area focused on advances in the efficiency of the algorithm on large-scale computing platforms. These advancements enabled us to significantly improve our effort in the CASP community-wide competition for protein structure prediction. During the CASP5 competition in summer of 2002, we attempted blind predictions of 20 proteins with sizes and structures far more complex than any previously attempted by our group. The results of assessment for the category of proteins which contain "new folds" (i.e. proteins which were not known by existing databases) show that our group's predictions were in the top 10% of groups which predicted these targets. Much of this research is joint work with the bio-chemistry group headed by Dr. Teresa Head-Gordon at UC Berkeley, and the computer science group of Dr. Silvia Crivelli at Lawrence Berkeley Laboratory.

2 Objectives of this Research

The primary objective of this research is to develop large-scale optimization methods for optimization problems that arise in molecular chemistry. The optimization approaches are being developed, however, in a manner that makes them applicable to a broad class of large-scale optimization problems. The main optimization problem that is being considered is the large-scale global optimization problem. This is the key problem that must be solved to determine the configuration of a molecule or macro-molecule once its potential energy function is known. Therefore the primary objective of this research is to develop efficient and effective large-scale global optimization methods for determining the structure of proteins. This includes testing of the methods on problems that constitute the state-of-the-art at the current time in optimization approaches to protein folding. Since the solution of large scale unconstrained optimization constitutes the major cost of

the global optimization calculations, another research emphasis is finding ways to reduce the costs of the local optimizations in the context of the global optimization algorithm. In addition, due to the size and difficulty of the problems being solved, their solution requires the use of very powerful computers, generally parallel computers. Therefore development of efficient parallel large-scale global optimization methods is an objective of this research.

In addition to the global optimization research, another important objective of this research includes the development of interior point methods for for very large equality and inequality constrained optimization. Both of these topics are of basic importance to the Air Force, with applications ranging from construction of new polymeric materials for the global optimization work, to a wide range of design, scheduling and control problems such as aircraft design and optimal trajectory problems for the constrained optimization research.The final thrust of this research has been in the area of large scale tensor methods for large-scale systems of nonlinear equations. This work has applications in the simulation of complex physical models and is also of major importance to the Air Force.

3 Global Optimization Methods for Protein Folding Problems

Working in conjunction with biochemists and computer scientists at U.C. Berkeley and Lawrence Berkeley Laboratory, we continue to be one of the leading groups that is taking an optimization-centered approach to one of the leading computational challenges, the solution of protein folding problems. In the past two years, we have expanded and enriched our group's approach so that it is now able to deal with proteins of the most general complexity (including alpha helices and beta sheets) and of as large a size as is being handled computationally (up to about 400 amino acids). Our group has competed with very good results in both of the last two CASP (Critical Assessment of Techniques for Protein Structure Prediction) competitions, in summer/fall 2000 and summer/fall 2002.

Our main research accomplishment has been the critical advancement in the ability of our algorithm to predict proteins of increasingly complex structure. Through the development and testing of our optimization-based approach for incorporating predicated beta-strand and beta-sheet structure, our method has acheived some notable computational success in predicting the structure of extremely difficult targets in the recent CASP5 competition.

In previous research periods, we had developed a stochastic/perturbation approach for solving global optimization problems from molecular chemistry. The key to this approach is a step that performs a small-scale global optimization with only a small number of parameters variable and the remaining parameters temporarily fixed, followed by a large-scale local minimization with all variables varying.

The stochastic/perturbation global optimization method applied to protein structure prediction consists of two phases. The first phase generates a set of good initial protein structures using information specific to the domain of protein structure prediction. This information is the prediction of secondary structure, which is readily available through

servers with accuracies averaging around 75%. Given the widespread use and high accuracy of secondary structure prediction methods, it would be impractical not to attempt to utilize them in this problem domain and only by using them can we realistically tackle reasonable-sized problems.

Secondary structure prediction is incorporated into the first phase with the goal that the second phase begins with structures that have much of the predicted secondary structure, but not tertiary strructure. In previous research periods, this was accomplished solely through the use of "biasing" functions, which are penalty terms that bias the structures towards predicted secondary structure, used within the context of local minimization. Recently, we have incorporated a powerful visualization tool, developed by our colleagues at Berkeley and LBL, to use in conjuction with the biasing functions to more efficiently create starting structures with predicted secondary structure. This technique was especially useful in creating starting structures for protein targets with extremely complex secondary structures, and without the visualization tool, it would not have been possible to predict structures for many of the difficult targets we attempted to predict in CASP5.

The second phase of the algorithm improves upon the initial structures by a combination of breadth (work on a variety of structures) and depth (improve the current lowest energy structures). It does this by a combination of steps that change the structure in a restricted manner – this is accomplished by small scale global optimization using a stochastic method applied to a selected small subset of the parameters – and local minimizations applied to the full structure. A key to this step is selecting a small subset of parameters that, by changing their value, can lead to a substantially improved structure for the entire problem. A list is kept of all minima obtained from the local minimizations, and after some (fixed) number of iterations, this list is clustered via pairwise root mean squared deviation (RMSD) evaluations, and ordered by energy value within each cluster. The RMSD of 2 configurations measures the closeness of their structures, i.e. a few Å difference means the structures are very similar, and beyond 15Å usually means they are not. After clustering, if the stopping criteria have not been met, additional iterations of the second phase are applied, starting from a new list containing the lowest energy minimizer from each cluster.

A very general framework that shows how the parts of the global optimization approach fit together is given below in Algorithm 1.

The second phase of the stochastic/perturbation algorithm requires many compute cycles to perform the small-scale global optimizations and full-dimensional local optimizations. During the period of this research grant, we have successfully adapted our algorithm to several massively parallel computing environments, including the Cray T3E at the Pittsburgh computing center and the IBM SP RS/6000 at NERSC. Computational requirements for the exremely large proteins tackled in CASP5 required the use of from 64 to 256 prosessors to complete the predictions within the submission deadlines for the competition.

Our use of smoothing within our global optimization algorithm has has been an important component over the past several years, and we present our current experimental findings here. By smoothing we mean replacing the objective (energy) function with a smoother version that is intended to retain the coarse grain structure of the original

Algorithm 1: Stochastic/Perturbation Global Optimization Algorithm for Protein Structure Prediction

1. **Phase 1: Coarse Identification of Configurations :**
   Generate initial configurations containing domain specific tendency information.
   (a) **Generate configurations with (server obtained) predicted secondary structure using the visualization tool**
   (b) **"Biased" Full-Dimensional Local Minimizations :**
   Use server obtained information to create biasing terms for predicted secondary structure. Perform a local minimization from a subset of the sample points, using "biasing" penalty functions to superimpose predicted secondary structure on the standard energy surface.
   (c) **"Unbiased" Full-Dimensional Local Minimizations :**
   Perform a local minimization from each "biased minimizer" using the "unbiased" energy function. Save these local minimizers for Step 2a.
2. **Phase 2: Improvement of Local Minimizers:**
   For some number of iterations:
   (a) **Select a Configuration and Small Subset of Parameters to Improve:**
   From the list of full-dimensional local minimizers, select a local minimizer to improve. Then select a subset of the parameters of this configuration to optimize in step 2b.
   (b) **Small Sub-problem Global Optimization:**
   Apply a global optimization algorithm to the energy of the selected configuration with only the selected parameters as variables.
   (c) **Full-Dimensional Local Minimization :**
   Apply a local minimization procedure, with all the parameters as variables, to the lowest energy configurations that resulted from step 2b, and merge the new local minimizers into the list of local minimizers.
   (d) **Cluster Local Mimima and Test for Convergence**
   Cluster the list of local minimizers via pairwise RMSD and if the stopping criteria has not been met, repeat all steps of Phase 2 from a new list of local minimizers containing the lowest energy minimizer from each cluster.

function but to damp the finer grain structure. The smoothed function should thus be easier to locate low minimizers on, while still serving as a guide to the structure of the original energy function. Previous to this research grant we had developed novel analytic techniques for smoothing molecular chemistry objective functions. They rely on replacing the Lennard-Jones and electrostatic terms with parameterized, smoother variants. Our research confirmed that our approaches definitely reduces the number of minimizers substantially. It also revealed the phenomenon of "order flips", that the lowest minimizer of the smoothed function may not track back to the lowest minimizer of the unsmoothed function. This indicated that global optimization techniques remain crucial even with smoothing.

Results on simple proteins, polyalanine and metenkephalin, showed that the use of smoothing significantly enhances the efficiency of our global optimization approach. Our more recent experimental work, particularly on some of the CASP5 targets subsequent to the competition, has once again shown the utility of using smoothing. The success of smoothing compared to no smoothing is particularly evident on alpha-helical proteins,

but not as obvious with beta proteins. We see good potential for improving this experimental work through incorporation of smoothing, and intend to generalize the success on helical proteins to all protein structural types.

Some of the our recent successes in protein structure prediction include: (1.) During the CASP4 (summer/fall 2000) competition, we submitted the best prediction of any of the roughly 150 groups on one of the large alpha-helical proteins, and our submissions were among the best predictions on all of the several "new folds" that we attempted. New folds, where the structure does not closely match the structure of known proteins, are the hardest problems and the ones for which optimization-centered approaches are primarily intended. (2.) In the CASP5 (summer/fall 2002) competition, our submissions were among the best predictions on the ten "new fold" targets that we attempted. Our predictions on these targets ranged from the 70th to the 98th percentile among all the 165 groups in the competition. Overall we placed about 15th of the 165 groups in the assessments of the new fold category; most of these groups are using far more chemical knowledge that our more purely optimization-centered approach. We appear to be one of the two leading groups (with roughly equal performance) whose protein structure predictions are based primarily on energy optimization.

4 Algorithms and Software for Large Scale Constrained Optimization

As part of this grant we have been deeply involved in development of methods for large scale nonlinear optimization. We have been active in this area on three fronts: enhancement of the interior point algorithm KNITRO, a mathematical investigation into failure of line-search algorithms for such problems, and beginning development of a new constrained optimization algorithm based on successive linear programming.

The algorithm KNITRO (formerly called NITRO) was developed by us under a previous AFOSR grant, in collaboration with Jorge Nocedal at Northwestern University. It uses trust regions in conjunction with an interior point strategy, to deal with large numbers of inequality constraints, by approaching the optimum from the interior of the feasible region. Over the last couple of years KNITRO has developed into one of the best performing codes for large-scale constrained optimization problems, together with LOQO, developed at Princeton and SNOPT, developed at Stanford and UCSD. This assessment is borne out in benchmarking performed by Hans Mittelmann at Arizona State University. In the period 2000-2003 we have concentrated on making KNITRO into a useful and available piece of software, and on some crucial extensions to its capabilities. In addition we completed a paper analyzing the convergence of the algorithm which appeared in *Mathematical Programming*.

In the last two years, KNITRO has been made available on the NEOS server. (NEOS is an on-line facility maintained by Argonne National Laboratory, that is available to the general public and solves over 10,000 user-submitted optimization problems a month.) In the first 4 months of this year NEOS users have accessed KNITRO over 2700 times. It has been used to solve problems in many applications areas. A group of us has formed a company, Ziena Optimization, which is marketing the most powerful version of KNITRO

to companies, and which has received an SBIR grant.

In the last two years we have made several important improvements to this algorithm. The original version of KNITRO used exact second derivatives to approximate the Lagrangian. We did have have a quasi-Newton version, but that approach is impractical for problems with many degrees of freedom. Now we have developed a finite-difference truncated Newton version of the algorithm. This established approach uses gradient information gathered in the course of a conjugate gradient iteration to develop an implicit approximation to the Hessian. It fits naturally into the KNITRO framework. This allows us to obtain fast convergence using only gradient information, and without constructing a large matrix to approximate the Hessian. Another improvement was to add a facility to detect when a problem had inconsistent or unsatisfiable constraints, and terminate the iteration with a message to that effect.

Although the iterates produced by KNITRO do not necessarily satisfy all inequalities at every iteration in the interest of efficience and flexibility, there are application problems where it is important for the iterates to be feasible in that sense. In order to develop a version of KNITRO with that feature we had to derive a new type of Cauchy step to handle the interaction of the trust region and inequality constraints. This version is now available, and we have described our technique in a paper that has appeared in the journal *Computational Optimization and Applications*.

We have also in the course of this research made a mathematical study of failures of line search methods, which is relevant to KNITRO. One of the features of KNITRO is its use of a trust region rather than a line search to control the step. The advantage of this is that we can obtain a more robust guarantee of global convergence. In order to assess the importance of this issue we made a study of the kinds of problem where a line search method would fail and a trust region would not. We were able to establish that whenever the constraint derivative matrix has a rank deficiency of order one at a point, there is about an even chance that a line search method will fail near that point. This is an important issue, there are known problems involving nonlinear partial differential equations where line search methods fail in practice, and our study indicates that it is worthwhile to try to extend trust region methods to these problems. As part of this study, we were also able to show that such failures were much less likely to occur in unconstrained optimization, and this may account for the similar performance of trust region methods and line search methods on such problems. A paper describing this work will appear in *Mathematical Programming* in early 2004.

A final aspect of our constrained optimization research involves beginning development of a quite different approach to nonlinearly constrained optimization, that should have crucial advantages in an important class of problems. A weakness of interior point methods like KNITRO is that they perform best when the starting point is far from all inequality constraints. This means that these methods tend to make poor use of a starting point close to the solution, Such a "warm start" is often available, for example, when solving a sequence of closely related problems.

To be able to handle very large instances of such problems we have developed an active set method using repeated solution of linear programs. Active set methods tend to approach the solution along the boundary of the feasible region, and are able to make better use of a warm start. However, existing active set methods, most of which use

successive quadratic programming, are limited in the size of problems they can solve, especially when the number of degrees of freedom is large. To be able to handle larger problems we have turned to successive linear programming as a means of estimating the correct active set. We then solve an equality constrained quadratic minimization problem, using this active set. This approach is related to an earlier method due to Roger Fletcher, but makes more use of trust regions, and is better able to handle negative curvature. Preliminary experiments with such a method show it to be efficient and quite robust compared to other methods. A paper describing this method has been accepted by *Mathematical Programming Series B*. Currently we are working on making this method more efficient by using inexact solutions of the linear programs.

5 Tensor Methods for Large Scale Systems of Nonlinear Equations

The main objective of this research was to combine multiple approaches for solving systems of nonlinear equations, primarily tensor methods and Krylov subspace methods, to develop an effective large-scale nonlinear equations solver. We developed three Krylov-based methods for iteratively solving the local tensor model and embedded them within an inexact framework for solving large, sparse nonlinear equations problems. Over the course of this research, we made several contributions, and we summarize the four principal ones here.

First, we investigated the performance of direct tensor methods on smaller problems over a range of conditionings, from well-conditioned to ill-conditioned to singular. Prior to this investigation, studies on direct tensor methods only focused on singular problems or on general problems that are well-conditioned, and had not investigated the performance of tensor methods on ill-conditioned problems. The results showed that direct tensor methods outperform Newton's method as the problems become more ill-conditioned. Specifically, the results show that, eventual quadratic convergence notwithstanding, the performance of Newton's method degrades as ill-conditioning grows, whereas tensor methods are relatively unaffected to only mildly affected (for larger rank deficiencies). This observation has particular relevance for practical problems where the Jacobian at the solution is very ill-conditioned, such as in the identification of bifurcation points or in path following problems that try to locate turning points.

Second, we developed a new robust global strategy for tensor methods, called the curvilinear linesearch. This linesearch combines the Newton and tensor steps into a single parametric step, which guarantees a monotonic decrease on the tensor model and also asymptotically approaches the Newton direction as the step length shrinks to zero, guaranteeing descent on the nonlinear equations. These theoretical properties make the curvilinear linesearch more attractive than other ad hoc linesearches that use the tensor step and Newton direction separately, such as the standard tensor and TENSOLVE linesearch. Our tests showed that the curvilinear linesearch outperforms the standard tensor and TENSOLVE linesearches, both on a small-scale test set and on large-scale problems. The curvilinear linesearch tensor method also appears to be more robust, in terms of success versus failure in solving problems, than both linesearch-based Newton's

method and the standard tensor linesearch and is roughly comparable to the TENSOLVE linesearch. Due to better theoretical properties, greater efficiency, and more robust performance, we believe that the curvilinear linesearch is a worthy improvement over the current linesearches used in tensor methods.

Third, we developed three large-scale methods for iteratively solving the local tensor model using Krylov subspace techniques, which we denote block-3, block-2, and block-2+. The three methods share the use of block Arnoldi concepts but are distinctly different in their approach to solving the tensor model. Two key features of these new methods are that their costs are similar to GMRES, requiring only one Jacobian-vector product at each iteration and $O(nm)$ additional arithmetic operations beyond GMRES per solve, and that the step also satisfies the tensor model to within a specified tolerance, making it possible to control the quality of the step. The nearest competing large-scale method for solving the local tensor model is used in the Feng-Pulliam tensor-GMRES method, which solves a modified local model with a projected tensor term and computes a tensor step with an unpredictable residual error.

The new tensor-Krylov methods proposed in this research solve the local tensor model in a novel fashion. Compared to the other two large-scale tensor methods (the method of Bouaricha, and the tensor-GMRES method of Pulliam and Feng), they are the only large-scale tensor methods that produce an approximate tensor step from the unprojected tensor model and that approach the minimum residual of the exact tensor step to within a specified accuracy. For instance, the Bouaricha method solves the unprojected tensor model, but the relationship between the residual of the tensor model and the GMRES tolerance used in the two linear solves is not known. Also, the Feng-Pulliam method solves a projected tensor model, which loses some information in the projection, and again lacks a direct relationship between the GMRES tolerance and the residual. The new tensor-Krylov methods also can utilize much of the technology developed for Newton-Krylov methods, including preconditioning and restarting.

Last, we incorporated these three local solvers into an inexact nonlinear solver framework with various linesearch options, including the curvilinear linesearch, for different versions of a "tensor-Krylov" method, which we denoted TK3, TK2, and TK2+. We implemented these methods in NOX, which is Sandia's production-quality code that hosts a suite of nonlinear solvers. We then compared the performance of the methods of Bouaricha, Feng and Pulliam, and our new methods on a set of large-scale fluid flow problems.

Among all of the methods, the new tensor-Krylov methods and the tensor-GMRES method by Feng and Pulliam are the most competitive. Our results suggest that these large-scale tensor methods hold advantages over Newton-GMRES, particularly on ill-conditioned problems. In general, the new methods have proven to be more robust than Newton-GMRES on most problems, especially when the tensor-Krylov methods employ the curvilinear linesearch.

For some problems, the projected tensor model is sufficient for a good step, which gives the advantage to tensor-GMRES because it has a more efficient local solver with the scalar implementation of GMRES. For other problems, the tensor-Krylov methods are more efficient, presumably because tensor steps from the projected model are inferior.

When considering the computational efficiency of the methods, Newton-GMRES al-

most always performs the best on the easier problems, but it performs poorly on the more difficult problems, where the tensor-Krylov methods are more efficient. Among just the new tensor-Krylov methods, it appears that TK2 and TK2+ edge out TK3 in terms of efficiency, particularly when using restarts. There is usually little difference between TK2 and TK2+, so the additional complexity of TK2+ may limit its usefulness in practice.

Thus, the new tensor-Krylov methods are especially aimed at large-scale problems that are highly ill-conditioned or singular, where algorithms based on Newton's method exhibit very slow convergence and where tensor methods excel. Of particular interest are PDE problems, which are the focus of many large-scale algorithms. For example, multigrid methods may solve a variety of PDE problems efficiently, but the coarse grid solution can be difficult to obtain in some cases, which could be a worthy application of tensor-Krylov methods. Our results suggest that the tensor-Krylov methods are, on average, more efficient and robust than Newton-GMRES. In addition, our results show that tensor-GMRES is a viable and competitive algorithm that should be considered.

6 Publications resulting from this grant

A. Azmi, R. Byrd, E. Eskow and R. Schnabel. New smoothing techniques for global optimization in solving for protein conformation, accepted for publication in *Global Optimization Case Studies*, J. Pinter, ed., Kluwer Academic Publishers.

A. Azmi, R. Byrd, E. Eskow, R. Schnabel, S. Crivelli, T. Philip and T. Head-Gordon(2000). Predicting Protein Tertiary Structure Using a Global Optimization Algorithm with Smoothing, in *Optimization in Computational Chemistry and Molecular Biology: Nonconvex Optimization and Its Applications*, C.A. Floudas and P.M. Pardalos, eds., Kluwer Academic Publishers, pp. 1–18.

Brett W. Bader (2003). Tensor-Krylov Methods for Solving Large-scale Systems of Nonlinear Equations, Ph.D. Thesis.

B. W. Bader and R. B. Schnabel. Curvilinear linesearch for tensor methods, to appear in *SIAM J. Sci. Comput.*

R.H. Byrd, J.C. Gilbert and J. Nocedal(2000). A trust region method based on interior point techniques for nonlinear programming, *Mathematical Programming* **89**, pp. 149–185.

R. Byrd, N. Gould, J. Nocedal, and R. Waltz. An Active-Set Algorithm for Nonlinear Programming Using Linear programming and Equality Constrained Subproblems, to appear in *Mathematical Programming series B.*

R. Byrd, N. Gould, J. Nocedal, and R. Waltz. On the convergence of successive linear programming algorithms, submitted to *SIAM Journal on Optimization.*

R. Byrd, M. Marazzi and J. Nocedal. On the Convergence of Newton Iterations to Non-Stationary Points, To appear, *Mathematical Programming series A* .

R. Byrd, J. Nocedal and R. Waltz (2003). Feasible Interior Methods Using Slacks for